

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/575,560A
Source: 1KWP
Date Processed by STIC: 11/3/06

ENTERED



IFWP

RAW SEQUENCE LISTING

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:26

Input Set : A:\33412_SEQ_ID_ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw

```

3 <110> APPLICANT: Sandoz GmbH
5 <120> TITLE OF INVENTION: Process for Producing Penicillin
7 <130> FILE REFERENCE: 33412US-PCT
9 <140> CURRENT APPLICATION NUMBER: US/10/575,560A
10 <141> CURRENT FILING DATE: 2006-04-11
12 <150> PRIOR APPLICATION NUMBER: PCT/EP2004/011566
13 <151> PRIOR FILING DATE: 2006-10-14
15 <160> NUMBER OF SEQ ID NOS: 8
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 411
21 <212> TYPE: PRT
22 <213> ORGANISM: Penicillium chrysogenum
24 <400> SEQUENCE: 1
26 Met Val Asp Pro Ser Val Ser Gly Ile Thr Lys Met Asp Thr Asn Asp
27 1 5 10 15
30 Ile Lys Gln Asn Asp Ile Pro Lys Asp Gln Pro Thr Leu Val Arg Trp
31 20 25 30
34 Tyr Met Asp Val Arg Arg Trp Asp Glu Lys Tyr Phe Asp Leu Pro Leu
35 35 40 45
38 Leu Glu Thr Leu Thr Gln Pro Asp Gln Ala Ala Val Lys Lys Tyr Tyr
39 50 55 60
42 Gln Thr Ser Asp Lys Arg Leu Ser Leu Ala Ser Gln Leu Leu Lys Tyr
43 65 70 75 80
46 Tyr Tyr Ile His Gln Ala Thr Gly Thr Pro Trp Ser Lys Ile Glu Ile
47 85 90 95
50 Gln Arg Thr Pro Met Pro Glu Asn Arg Pro Phe Tyr Asp Ser Ser Leu
51 100 105 110
54 Asp Phe Asn Val Ser His Gln Ala Gly Leu Thr Leu Phe Ala Gly Thr
55 115 120 125
58 Arg Ala Ala Thr Ala His Ser Leu Ser Gly Gly Pro Gln Thr Leu Pro
59 130 135 140
62 Arg Val Gly Ile Asp Val Ala Cys Val Asp Glu Pro Ser Arg Arg Arg
63 145 150 155 160
66 Ala Asn Arg Pro Pro Lys Thr Leu Ala Asp Leu Ala Thr Phe Val Asp
67 165 170 175
70 Val Phe Ser Asp Val Leu Ser Leu Arg Glu Leu Ala Thr Ile Lys Asn
71 180 185 190
74 Pro Tyr Ala Thr Leu Lys Leu Ala Arg Glu Leu Gly Leu Asn Lys Ser
75 195 200 205
78 Asp Pro Ser Lys Asp Asp Gln Glu Val Leu Ala Ala Tyr Gly Ile Arg
79 210 215 220
82 Leu Phe Tyr Ser Ile Trp Ala Leu Lys Glu Ala Tyr Leu Lys Met Thr

```

RAW SEQUENCE LISTING

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:26

Input Set : A:\33412_SEQ ID.ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw

```

83 225          230          235          240
86 Gly Asp Gly Leu Leu Ala Ser Trp Ile Lys Asp Leu Glu Phe Thr Asn
87          245          250          255
90 Val Val Pro Pro Glu Pro Val Gln Thr Val Gly Phe Ala Gly Asp Pro
91          260          265          270
94 Ser Ala Thr His Ala Pro Ser Val Gln Asn Trp Gly Arg Pro Tyr Ser
95          275          280          285
98 Asp Val Lys Ile Ser Leu Arg Gly Ile Pro Asp His Ser Val Arg Val
99          290          295          300
102 Gln Pro Val Gly Phe Glu Ser Asp Tyr Ile Val Ala Thr Ala Ala Ser
103 305          310          315          320
106 Gly Pro Asn Ile Gly Ser Val Ser Arg Gln Val Val Val Asn Asp Ser
107          325          330          335
110 Asp His His Leu Pro Gly Arg Ile Thr Ala Phe Asp Ser Glu Thr Gly
111          340          345          350
114 Leu Gln Asn Val Arg Ile Pro Pro Ile Ala Leu Arg Ser Ile Gly Asp
115          355          360          365
118 Gly Asp Pro Trp Arg Val Asp Ser Lys Ile Ser Asp Pro Trp Leu Pro
119          370          375          380
122 Met Gln Glu Val Asp Ile Glu Ile Asp Ile Arg Pro Cys Ala Asp Gly
123 385          390          395          400
126 Arg Cys Glu His Leu Arg Asp Leu Pro Ser Phe
127          405          410
130 <210> SEQ ID NO: 2
131 <211> LENGTH: 1284
132 <212> TYPE: DNA
133 <213> ORGANISM: Penicillium chrysogenum
135 <400> SEQUENCE: 2
136 atggttagacc ccagtgtgtc tggaattgtg agtagccaca tagcctccat gagtgcaccc 60
138 actgaccaat ttcagaccaa aatggatacc aatgatatca aacagaatga catccccaag 120
140 gaccagccca cggttggtccg atggtacatg gatgtcagac gttgggatga aaaatacttt 180
142 gatctccctt tgcttgaaac cttaacacag cctgatcagg cagctgtcaa gaagtactat 240
144 caaacatcgg acaagcgcct gtccttggtc tcccagttgc tgaaatatta ctacattcac 300
146 caagccactg gcactccctg gagcaagatt gagatccagc gtactccgat gcccgaat 360
148 cgaccattct acgattcaag cctggatttc aacgtcagcc atcaggctgg tctcactctg 420
150 ttcgcaggca cgcgtgccgc aacagcccac tccttatccg gtggacctca aacattgcct 480
152 cgcgtgggaa ttgacgttgc gtgtgttgat gaacctctc gtcgtcgtgc taatcgtccc 540
154 ccgaagacac ttgccgacct tgcaaccttc gtggatgtct tcagtacgt tctctcactc 600
156 cgtgagcttg cgaccatcaa gaacccttac gcgactctta aattggctcg tgagcttggt 660
158 ctgaataaaa gtgacctgag caaagacgac caggaagtcc ttgctgccta cggcattcgg 720
160 ctgttctact cgatttgggc tctcaaggag gcttacttga aaatgaccgg agacggcctt 780
162 ctggcctctt ggataaagga tctggaattc acaaacgttg ttcccccgga accagttcaa 840
164 acagtcggat ttgctggtga tccttctgcc actcacgccc cctcggtcca aaattggggc 900
166 cggccttact ccgatgtcaa aatctccttg cgtggcattc ctgaccattc tgtgcgcgtt 960
168 cagctcgtcg gcttcgagtc cgactacata gttgccacgg ccgcgtcggg cccaatatt 1020
170 ggatccgttt cgcggcaggt agtcgtgaat gacagcgatc accatctgcc agggcgtatc 1080
172 acagccttcg actctgagac tggactccag aacgtccgca ttcccccaat cgcgcttcga 1140
174 tcaattggcg atggggaccc ctggcgtgtg gactcgaaaa tcagcgaccc ctggctcccc 1200
176 atgcaggagg tcgatattga aatcgatatc cggccctgtg cggatggctg ttgcgagcac 1260

```

RAW SEQUENCE LISTING

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:26

Input Set : A:\33412_SEQ ID.ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw

```

178 ctacgggatt taccaagctt ttaa                                     1284
181 <210> SEQ ID NO: 3
182 <211> LENGTH: 1236
183 <212> TYPE: DNA
184 <213> ORGANISM: Penicillium chrysogenum
186 <400> SEQUENCE: 3
187 atggttagacc ccagtgtgtc tgggaattacc aaaatggata ccaatgatat caaacagaat      60
189 gacatcccca aggaccagcc cacgttggtc cgatggtaca tggatgtcag acgttgggat      120
191 gaaaaataact ttgatctccc tttgcttgaa accttaacac agcctgatca ggcagctgtc      180
193 aagaagtact atcaaacatc ggacaagcgc ctgtccttgg cctcccagtt gctgaaatat      240
195 tactacattc accaagccac tggcactccc tggagcaaga ttgagatcca gcgtactccg      300
197 atgcccgaaa atcgaccatt ctacgattca agcctggatt tcaacgtcag ccatcaggct      360
199 ggtctcactc tgttcgcagg cacgcgtgcc gcaacagccc actccttatc cggtgggacct      420
201 caaacattgc ctgcgtggg aattgacgtt gcgtgtgttg atgaaccctc tcgtcgtcgt      480
203 gctaactcgt ccccgagac acttgccgac cttgcaacct tcgtggatgt cttcagtgac      540
205 gttctctcac tccgtgagct tgcgaccatc aagaaccctg acgcgactct taaattggct      600
207 cgtgagcttg gtctgaataa aagtgacccg agcaaagacg accaggaagt ccttgctgcc      660
209 tacggcattc ggctgttcta ctcgatttgg gctctcaagg aggccttactt gaaaaatgacc      720
211 ggagacggcc tttctggcctc ttggataaag gatctggaat tcacaaacgt tgttcccccc      780
213 gaaccagttc aaacagtcgg atttgcttggt gatccttctg ccactcacgc gccctcggtc      840
215 caaaattggg gccggcctta ctccgatgtc aaaatctcct tgcgtggcat tcctgaccat      900
217 tctgtgcgcg ttcagcccgt cggcttcgag tccgactaca tagttgccac ggccgcgtcg      960
219 ggccccaaata ttggatccgt ttcgcggcag gtagtcgtga atgacagcga tcaccatctg      1020
221 ccagggcgta tcacagcctt cgactctgag actggactcc agaacgtccg cattccccca      1080
223 atcgcgcttc gatcaattgg cgatggggac ccctggcgtg tggactcgaa aatcagcgac      1140
225 ccctggctcc ccatgcagga ggtcgatatt gaaatcgata tccggccctg tgcggatggt      1200
227 cgttgcgagc acctacggga tttaccaagc ttttaa                                     1236
230 <210> SEQ ID NO: 4
231 <211> LENGTH: 3188
232 <212> TYPE: DNA
233 <213> ORGANISM: Penicillium chrysogenum
235 <400> SEQUENCE: 4
236 gtcgaccgaa gtggtttctcg ttcactcgca catcaagacc accgatcagc tcttgcccgc      60
238 ccttctttgt cttgttggca gactcggcaa gcaaaatgag cccggcgcat gtacccacg      120
240 tcggtttgcg atccactctg cataaccac gtattagatc gaattgatat ggactaacc      180
242 ggttcactca ctttacgaat tctcgcagt gctcgagaag atttgacctt gctgcgacta      240
244 aagacatagt ggtactctcg cctccgggca agaccaggcc gtcgcatgtt gccagtctt      300
246 gtggcgctcc tacttcaatg aagtgccatt ccgacggctg cgcttgctca gcggccttt      360
248 tcaaaagctg cacatgctca aagaatgcgc cctgtagggc caggactcca acagtgatag      420
250 ccatttcctc tgaagatcgg aattgcggac cctccgagct cgggtgcttc ttgatattga      480
252 tgactctttt taaagcacat gactttgact ttccggcggg gaacgtatca acacgtgatg      540
254 gcggcttatc tccatcttta attccacgcg acatcaggat atcgtgagag ctctcggacg      600
256 attcctgcgc actttgaaaa cagactgcat aaccgaggca ttatagtata aaacaaatag      660
258 actcacctac agaaagagtg ataagttagg tcctatacct gtttccaatg tttctctctc      720
260 ttgctggatc agctttaaca tatctatgga tggatcttg gatagtcata gtcattatgc      780
262 gcttgctatt gcatgtctct ttgctacatc ctatttatgg tattatgtac acggcctgtt      840
264 tctcgtttgc cggcctattg atgtatacat gtattggtgt aggtagtat tgccctgcct      900
266 tatcgacacg tgctgataga taaggacccc gataagacgc caacatggct tctatccagg      960
268 tgtggatgct ccgcatccaa ggtgcgaata tacgagatca caatgcaatg gtagacccca      1020

```

RAW SEQUENCE LISTING

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:26

Input Set : A:\33412_SEQ ID.ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw

```

270 gtgtgtctgg aattgtgagt agccacatag cctccatgag tgcacccact gaccaatttc 1080
272 agacccaaat ggataccaat gatatacaac agaatacat ccccaaggac cagccacagt 1140
274 tgggtccgat gtacatggat gtcagacgtt gggatgaaaa atactttgat ctccctttgc 1200
276 ttgaaacctt aacacagcct gatcaggcag ctgtcaagaa gtactatcaa acatcggaca 1260
278 agcgctgtc cttggcctcc cagttgctga aatattacta cattcaccaa gccactggca 1320
280 ctccctggag caagattgag atccagcgta ctccgatgcc cgaaaatcga ccattctacg 1380
282 attcaaacct ggatttcaac gtcagccatc aggtgtgtct cactctgttc gcaggcagc 1440
284 gtgccgcaac agccactcc ttatccggtg gacctcaaac attgcctcgc gtgggaattg 1500
286 acgttgctg tggtgatgaa ccctctcgtc gtcgtgctaa tcgtcccccg aagacacttg 1560
288 ccgaccttgc aaccttcgtg gatgtcttca gtgacgttct ctactccgt gagcttgcca 1620
290 ccatcaagaa ccgtacgcg actcttaaat tggctcgtga gcttggctg aataaaagt 1680
292 acccgagcaa agacgaccag gaagtccttg ctgacctagg cattcggtg ttctactcga 1740
294 tttgggctct caaggaggct tacttgaaaa tgaccggaga cggccttctg gcctcttgga 1800
296 taaagatct ggaattcaca aacgttggtc cccccgaacc agttcaaaaca gtcggatttg 1860
298 ctggtgatcc ttctgccact cagcgccct cggtccaaaa ttggggccgg ccttactccg 1920
300 atgtcaaaat ctcttgcgt ggcatctcgt accattctgt gcgcgttcag ctgcgcgt 1980
302 tcgagtcgga ctacatagtt gccacggcgc cgtcgggccc caatatgga tccgtttcgc 2040
304 ggcaggtagt cgtgaatgac agcgatcacc atctgccagg gcgtatcaca gccttcgact 2100
306 ctgagactg actccagaac gtccgcattc ccccaatcgc gcttcgatca attggcgatg 2160
308 gggacccctg gcgtgtggac tcgaaaatca gcgacccctg gctccccatg caggaggctc 2220
310 atattgaaat cgatatccgg ccctgtgcgg atggtcgtt cgagcaccta cgggatttac 2280
312 caagctttta aattccttct tgctgggata tgaccaggcg accatgcacc cgagttattt 2340
314 gcatattgca tctcctcatc tcatattcct ttctgagcgt gtttttcgga gcgataatta 2400
316 cccttgaaca tatttctgca ttgctgtatt gccattagcg aaaattccc agctagtgtg 2460
318 agttgatttc ctggaacgct gggggagtg cgtcagatg ttcatctcca ataagcccct 2520
320 caatgaatct tcacttcac ggatccaagg tcaatcttc agatcaagt caagtgtccc 2580
322 agaaagcacg ggtaaagaaa ccaagcctat ttctattcta tggctaatg taaactaaaa 2640
324 atgtagaagg aagaaaagca agtatccaac agtaggcggg tcatgacat cgtgtgcgct 2700
326 aaggatatat acatttcgaa ttgcaaagag ggaagagggt aatcaggagt gaaatgtgtg 2760
328 tcaagaggca atgtcaatgt caagatcatt gttgctctca tgagcagtca cggatttgt 2820
330 cggattgttc ggcgtctggg gccctcagat tctatttctg ggtcatgagc ttgagagtag 2880
332 gtaccgaaga agtgagcagt attatactgc agtgagtgt tagggggaat tccttctggt 2940
334 gaattgtggc gttcggggtt gctctccggt cttatgggtc ttaatctgga tgcccgatag 3000
336 tgcacccaag ttaggagaaa aacatatggt aagtgttaat cgtggagcag tgtggcgaat 3060
338 cgcgaattgg gtttggcact tagatttcca tggcgctaga gacgccgttg gcgcgagcac 3120
340 catcgacctc atttttatgc gcgtgggaca ttgctgcaag agttttgagc atcgaatccc 3180
342 gcgtcgac

```

345 <210> SEQ ID NO: 5

346 <211> LENGTH: 37

347 <212> TYPE: DNA

348 <213> ORGANISM: Artificial Sequence

350 <220> FEATURE:

351 <223> OTHER INFORMATION: oligonucleotide primer

353 <400> SEQUENCE: 5

354 ccccgctcgac cgaagtgggt tcggttcact cgcacat

37

357 <210> SEQ ID NO: 6

358 <211> LENGTH: 37

359 <212> TYPE: DNA

360 <213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:26

Input Set : A:\33412_SEQ ID.ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw

362 <220> FEATURE:

363 <223> OTHER INFORMATION: oligonucleotide primer

365 <400> SEQUENCE: 6

366 ccccgctcgac gcgggattcg atgctcaaaa ctcttgc

37

369 <210> SEQ ID NO: 7

370 <211> LENGTH: 41

371 <212> TYPE: DNA

372 <213> ORGANISM: Artificial Sequence

374 <220> FEATURE:

375 <223> OTHER INFORMATION: oligonucleotide primer

377 <400> SEQUENCE: 7

378 gggggccgag gcggcccatg gataccaatg atatcaaaca g

41

381 <210> SEQ ID NO: 8

382 <211> LENGTH: 43

383 <212> TYPE: DNA

384 <213> ORGANISM: Artificial Sequence

386 <220> FEATURE:

387 <223> OTHER INFORMATION: oligonucleotide primer

389 <400> SEQUENCE: 8

390 gggggccatt atggcctcat tcaggactac ctgccgcgaa acg

43

VERIFICATION SUMMARY

DATE: 11/03/2006

PATENT APPLICATION: US/10/575,560A

TIME: 10:30:27

Input Set : A:\33412_SEQ ID.ST25.txt

Output Set: N:\CRF4\11032006\J575560A.raw